

KAMİLƏ VƏLİYEVƏ (Azərbaycan)*
GÜLNARA CƏFƏROVA (Azərbaycan)**

MAŞIN TƏRCÜMƏSİNDE SÖZÜN SEMANTİK TƏHLİLİ

Xülasə

Dilçilikdə semantika mənəni və ya insan düşüncəsindəki konseptdə dil komponentlərinin (sözlərin, frazaların) bir-birinə necə uyğunlaşdığını öyrənir. Maşın tərcüməsində isə semantik təhlilin əsas prinsipi sözün mənasının aşkar çıxarılmasını öyrənməkdir. Bundan əlavə, semantik təhlilə aid sual-cavab və informasiyanın çıxarılması kimi müxtəlif praktik və nəzəri problemlər burada aşkarlanır.

Maşın tərcüməsinin semantik təhlil mərhələsində sözlərin mənasından çıxış edərək formal semantik modellər hazırlanır. Bu modellər sözün sintaktik və semantik təhlilinə təkan verir və semantik təhlil prosesinə kömək edir. Məqalədə maşın tərcüməsində sözün semantik təhlili məsələləri araşdırılır. Müəlliflər maşın tərcüməsində çoxmənalılıq probleminin çoxmənalı sözün mənasının yalnız kontekstdə müəyyənləşdirilməsi və yerində işlənməsini həyata keçirən program qurmaqla həll edilməsinin mümkünlüyünü göstərirlər.

Açar sözlər: *maşın tərcüməsi, semantik təhlil, söz, məna, model, lügət.*

Maşın tərcüməsi sistemlərinin icmali onu göstərir ki, ən az işlənmiş mərhələ semantik təhlil mərhələsidir. Bu onunla bağlıdır ki, sintaktik təhlildən fərqli olaraq semantik təhlildə söz formalarının mənalarının formalizə edilməsində ciddi çətinliklər vardır. Maşın insandan fərqli olaraq sözün məna çalarlarını kontekstə görə müəyyənləşdirə bilmir, çünki onda intuisiya, psixoloji hissələr kimi fizioloji keyfiyyətlər yoxdur. Semantik mərhələnin əsas məqsədi mətndə cümlələrin tərkibindəki sözlərin bir-biri ilə məna əlaqələrini təhlil etməkdir. Bunun üçün mütləq morfoloji və sintaktik təhlilə müraciət olunur. Semantik təhlildə maşın tərcüməsi məsələsinin həlli ilə əlaqədar biz təkcə ayrı-ayrı söz formalarının analiz və tərcüməsi ilə kifayətlənməmiş, həm də mətndə sabit söz birləşmələrinin analizi və tərcüməsinin seçilməsi prosedurunun düzgünlüyünü araşdırılmışdır. Məhz semantik mərhələ tərcümə vahidinin seçilmesinə, mətn və onun terminlərlə təsvirinin tamlığı və adekvatlığına, ifadənin, daha dəqiq desək, kiçik anlayış vahidlərindən təşkil olunmuş böyük vahidlərin anlaşılmasına cavabdehdir. Təsvir tipindən asılı olaraq tərcümənin kiçik

* Azərbaycan Milli Elmlər Akademiyası (AMEA), Nəsimi adına Dilçilik İnstitutu, Kompüter dilçiliyi şöbəsinin müdürü, dosent. E-mail:kamilavaliev@gmail.com

** Azərbaycan Milli Elmlər Akademiyası (AMEA) Nəsimi adına Dilçilik İnstitutu, Kompüter dilçiliyi şöbəsinin elmi işçisi, doktorant. E-mail:gulnarajafarova@yahoo.com

və böyük vahidləri kimi adətən bunlar təklif olunur: sözün özü, söz birləşmələri, mübtəda, xəbər, tamamlıq, zərflik qrupları, cümlə, giriş mətnində müntəzəm şəkillədə müşahidə edilən fragmənlər (onların tərcümə yaddaşında adekvat və normativ tərcümələri verilir) və s. (Vəliyeva, 2003: 38).

Semantik təhlil dildə ifadə vasitələrinin mənasının təcəssüm olunma bacarığını göstərən formal ifadə vasitələrinin və dildə ifadə olunan sözlərin uyğun mənasının göstərilməsi bacarığına malik alqoritmlərin öyrənilməsindən ibarətdir. Burada əsas məsələ sözün mənasının başqa sözlər, frazalar və morfemlərlə necə əlaqələndiyini aşkara çıxarmaqdır.

Q.Kazimova görə, semantik təhlil dilin məna zənginliklərini, sözün məna incəlik və çalarlarını, sənətkarın sözün semantik cəhətindən istifadə ustalığını öyrənmək üçün mühüm tədqiqat üsullarındandır. Belə təhlil vasitəsilə sözün daxili məzmunu, mahiyyəti açılır, yaxın mənalı sözlərin əlavə məna rəngləri, məzmun qohumluğu və məna əlaqəsi meydana çıxır (Kazimov, 2008: 145).

“Semantik təhlil” ifadəsi ilə nəyin nəzərdə tutulduğunu aydınlaşdırır. T.İzutsunun fikrincə, semantik təhlil təkcə sözlərin mənalarını təhlil etməkdən ibarət deyildir, o həmçinin dünya, ətraf aləm haqqında anlayış və düşüncələri, dili istifadə edən xalqın düşüncəsini eks etdirir. Semantika bir millətin tarixinin bu və ya digər mühüm dövründə dünyagörüşünün mahiyyəti və quruluşu haqqında fəaliyyətdir (İzutsu, 1975: 15).

Semantik təhlil müxtəlif istiqamətlərdə aparılır. Linqvistikada o, linqvistik biliyin, linqvistik mənanın müəyyən edilməsi (mənanın qurulması), dilin emal edilməsindən (məlumatın çıxarılması və modellərin qurulması) ibarətdir. Semantik təhlil paraqraf, cümlə, fraza, ifadə səviyyələrində tutmuş bütövlükdə yazının məhələlərini əhatə edən sintaktik strukturla bağlı olan təhlil mərhələsidir (13).

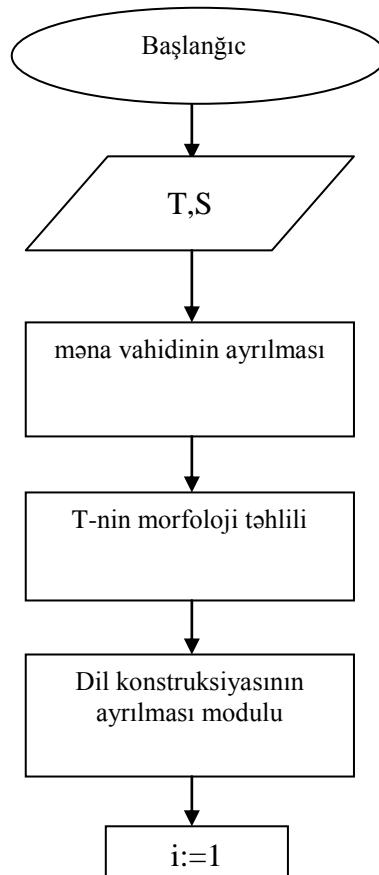
Mətnin avtomatik işlənmə sistemlərində semantik təhlil problemlərində bəhs edərkən B.Y.Qorodetski semantik inventarlaşdırma məsələləri üzərində dayanır. O, semantik inventarlaşdırmanı leksik-semantik və morfo-semantik səviyyələrdə aparır. Göstərilən bu səviyyələrdə paradiqmatik və sintaqmatik aspektlər mütləq nəzərə alınır. Semantik inventarlaşdırma prosesində söz və morfemlərin, onların müəyyən birləşmələrinin tezauruslarının tərtibi mühümdür (Vəliyeva, 2003: 16). Semantik təhlil üçün genişölçülü mətn korpusu və ya başqa növ korpusların tərtibi də mühüm amildir. Mətnin semantik təhlili statistik maşın tərcüməsi sahəsində xüsusi rol oynayır. Belə təhlil üçün ən vacib işlərdən biri ilk növbədə ümumi semantik lügətlərin və yuxarıda qeyd olunduğu kimi, tezaurusların tərtib edilməsidir. Onun ardınca semantik əlaqələrin təşkil edilməsi gəlir (Сокирко, 2001: 45). Semantik əlaqələrin aşkara çıxarılmasının ən effektli metodlarından biri leksik modellərin qurulması metodudur (Lyons, 1968: 428; Мочалова, 2014: 128-129). M.Hörsta görə, leksik modellərin qurulması metodu cins-növ əlaqələrinin müəyyənləşdirilməsi üçün tam uyğundur (Hearst, 1992: 540).

Maşında semantik təhlil mərhələsinə qədər morfoloji (Махмудов, 1982: 150) və sintaktik təhlil (Vəliyeva, 1996: 158) mərhələləri tamamlanmalı və düzgün alqoritmlər qurulmalıdır. Semantik təhlil alqoritminin qurulması təhlilin həllinin

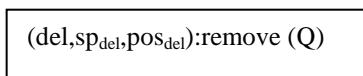
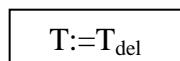
yarısı deməkdir. Rus tədqiqatçısı A.V.Moçalova alqoritmik təsvirdə istifadə edilmiş şərti işarələr vasitəsilə semantik təhlilin aparılması qaydasına əsaslanan sxem qurmuşdur (I sxem) (Мочалова, 2014: 128-129):

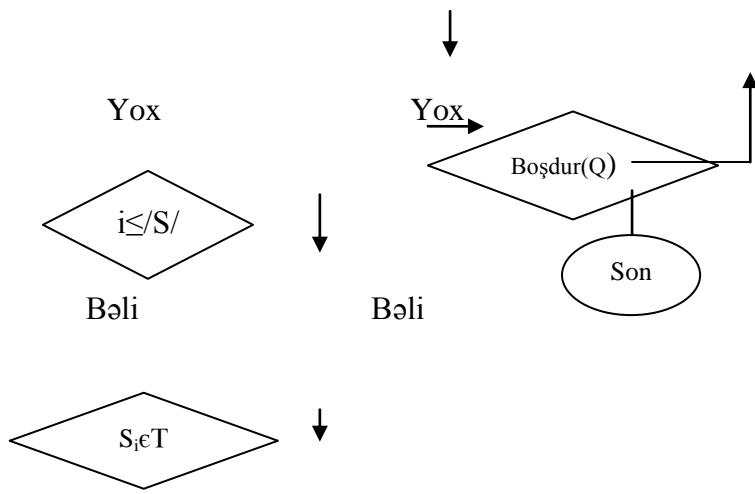
I sxem

Alqoritmik təsvirdə istifadə edilmiş şərti işarələr vasitəsilə semantik təhlilin aparılması qaydasına əsaslanan sxem



Yox ← ←





Bütün fərqli T çoxluqları üçün C_i uyğun gəlməli, yerinə yetirilməlidir.

daxil etmək ($Q, (t_{i2}, sp_i, pos(t_{i2}))$), t_{i2} – semantik tabelilik hissəsinin doğruluğu $R_i(t_{i1}, t_{i2})$, T -dən tapmaq

$R_i(t_{i1}, t_{i2}) \in BD$ -dir, o halda $R_i(t_{i1}, t_{i2})$ BD-yə semantik tabeliliyi əlavə etmək

T – təhlil üçün verilmiş cümlələr çoxluğu;

T_i – i – burada “ i ” bölünməz məna vahidi kimi təhlil üçün verilmiş “ T ” cümlələr çoxluğununu təklif edir;

S – əsas semantik şablonlar qrupu;

s_i – i – burada “ s ” əsas semantik şablonlar qrupu nümunəsidir;

sp_i – s_i – şablonu prioritet ardıcılılıq nümunəsidir;

$R_i(t_{i1} və t_{i2})$ – R_i semantik əlaqəni göstərir, s_i şablonunu müəyyənləşdirir və $t_{i1} və t_{i2}$ bölünməz məna vahidlərində əlaqə yaradır;

$pos(t_i)$ – t_i -dən çıxarılmış sonuncu sözün təhlil edilmiş cümlədə mövqeyi;

Q – ardıcılılıq növbəsi;

(t_{i2}, sp_i, pos_i) – Q ardıcılılıq növbəsinin elementi, s_i şablonu vasitəsilə verilmiş element.

Sözlər və cümlələrin quruluşu, sözlərin, frazeoloji ifadələrin, mətnlərin mənaları haqqında məlumat, onların məqsəd və nəticələri bu mərhələdə nəzərdə tutulur. Maşın tərcüməsinin bu mərhələsində sözlərin, söz birləşmələrinin və cümlənin mənalarının açılması mühüm əhəmiyyət kəsb edir. Bu mərhələdə X dilinin bir ifadəsinin Y dilindəki mənasının təmin edilməsinə cəhd edilir. Hesablama nöqtəyi-nəzərindən adekvat sayılan əsas həlletmə məsələləri bu sahə üçün təklif edilir.

Rus tədqiqatçısı A.Sokirkaya rusdilli mətnlər üzərində ilkin semantik təhlil məsələsini həll etməyə çalışmış, rus dilinin ümumi semantik lüğəti, semantik əlaqə, leksikanın ümumi kateqoriyaları, semantik xüsusiyyətlər, semantik xüsusiyyətlərin əlaqələri, DAİLİNQ proyektində söz birləşmələri, zaman qruplu lüğətlər (TimeRoss), tezaurus lüğətləri, lokativ tezauruslar (yer bildirən), maliyyə tezaurusları, kompüter tezaurusları, sabit söz birləşmələri lüğəti məsələlərini geniş araşdırılmışdır. O, programda semantik təhlilə giriş, semantik təhlildən çıxış, səthi semantik təhlil alqoritmi tərtib etmişdir. Onun fikrincə, semantik təhlil dildə işlənən cümlənin semantik strukturunu qurur. Semantik struktur semantik əlaqələri və semantik qovşaqları ifadə edir. A.Sokirkaya görə, semantik vahid semantik nəzəriyyənin mərkəzi konseptlərindən biridir. O, semantik bazanın əsasını təşkil edir (Сокирко, 2001: 120).

A.Sokirkanın fikrincə, semantik vahid mətndə aşkarla çıxan bütün valentlikləri ümumi şəkildə əhatə edir və daxildən öz konseptuallığını genişmiqyaslı linquistik mənbələrdən qoruyub saxlayır. Semantik vahid yalnız semantik təhlilin sonunda qurula bilər. Semantik təhlilin əsas məqsədi bütün valentliklərin qurulmasını təmin edən semantik vahidlərin konstruksiyasıdır. Həmçinin semantik mexanizmə daxil olan atributların və semantik vahidlərin semantik təhlili burada yerinə yetirilir. Əsas məsələ onların birini digərindən necə fərqləndirməkdir.

Təhlilin bütün mərhələlərində olduğu kimi, semantik komponentlər sözlərin orijinal variantından formallaşır. Semantik qovşaq nəzəriyyəsinin əsas mənbəyi təhlilin həyata keçirilməsidir. Əksər sintaktik qruplar semantik vahidlərə daxil ola, yaxud qovşaq atributlarına aid edilə bilər. Bir cümlənin semantik təsvirini vermək üçün semantik lüğətlər, tezaurus lüğətləri tərtib edilməlidir. Semantik quruluş sintaktik təsvirlərin daxil edilməsinə əsaslanır. Sözlərdən əlavə semantik vahidlərə söz birləşmələri, durğu işarələri də daxil edilir.

Semantik qovşağın əsas funksiyası aşağıdakılardır:

1. Atributlar dəstini təyin edir;
2. Qrafemlər dəstini təşkil edir;
3. Qovşaqdakı əsas semantik söz sayını təyin edir;
4. Qrammatik təhlili (daxili sintaktik xüsusiyyətləri) müəyyən edir;
5. Qovşaga aid olan fragmentlərin sayını təyin edir;
6. Qovşaq vasitəsilə idarə edilən sintaktik təhlili müəyyən edir;
7. Sözün tərkibindəki saysız-hesabsız elementlər və frazalar lüğətinə keçidi müəyyənləşdirir.

Semantik əlaqələr isə aşağıdakı parametrlərə malikdir:

1. Lüğətə girişə istinadı təyin edir;

2. Mətn daxilində mövcud olan sözlərin siyahısını müəyyənləşdirir;
3. Semantik əlaqələri asanlaşdırın sintaktik əlaqələri təyin edir;
4. Cümələ və ya ifadənin şərti olaraq uzunluğunu müəyyənləşdirir.

A.Sokırka semantik təhlil üçün belə bir sxem qurmuşdur (12) (II sxem):

II sxem

Semantik təhlil üçün ümumi sxem

1) Mətndən cümlənin çıxarılması

12) Tezaurus lüğət formalarının və quruluşlarının şərhi

2) Cümlənin morfoloji təhlilinə keçid

13) Leksik variantlar üçün ağacvari strukturun qurulması

3) Cümlənin sintaktik təhlilinə keçid

14) Semantik xüsusiyyətlərin yoxlanması

4) Semantik vahidlərin seçilməsi

15) Layihələndirmənin yoxlanması

5) Söz və ya ifadənin dəst halında qurulması

16) Səhv əlaqə konstruksiyalarının yoxlanması

6) Zaman qruplarının qurulması

17) Cümlələrarası əlaqələrin qurulması

7) Durğu işarələrində ifadə vasitələrinin qurulması

18) Anaforik əlaqələrin qurulması

8) Leksik vahidlərin təyini və qruplaşdırılması

19) Qiymətləndirmə

9) Sabit söz birləşmələrinin qruplaşdırılması

10) Leksik funksiyaların qurulması

11) Yerli qovşaqlar arasında əlaqələrin qurulması

Semantik təhlildə semantik əlaqələrin yaradılması üçün əvvəlcə şablonlar qurulur. Şablonlar məlumatların toplanmasında axtarış sisteminin bir hissəsidir. Burada əsas semantik modellərlə mətnin uyğunlaşdırılması təklif edilir. Bu isə təhlil edilmiş cümlənin semantik bağlılığı görə formalizə edilməsi ilə nəticələnir. Müəyyən şərtlər altında daha geniş təhlil təklifləri irəli sürür. O, əsas semantik şablonlarda təsvir edilir. Təbii dildəki semantik təhlilin təklif edilmiş metodu əsas semantik şablonların formalizəsini əhatə edir, lakin belə metodların sayı şablonların sayından daha azdır. Əsas semantik şablonların sayının az olması ilə əlaqədar olaraq bu istiqamətdə işlər sürətləndirilməlidir (Мочалова, 2014: 128-129).

Semantik təhlilin nəzəri problemləri çoxdur. Praktik cəhətdən də aşağı səviyəli emal mərhələlərində səhv'lərə səbəb olan problemlər müşahidə olunur. Məsələn, mütəxəssislər maşın tərcüməsində semantik təhlil zamanı problemlərin hansı xətt üzrə baş verdiyini müəyyənləşdirərkən belə məsələləri semantik problemlərdən daha çox maşın tərcüməsinin problemləri kimi qeyd edirlər. İki problem daha çox çətinlik törədir:

1) təhlil zamanı grammatik cəhətdən bir-biri ilə bağlı olan sözlər cütünün əlaqəliliyi, yəni emal zamanı mənanın qorunub saxlanması;

2) bir ifadənin bir neçə mümkün şərhlərindən hansının daha uyğun olduğunu müəyyənləşdirmək, yəni semantik ikimənalılığın aşkarılması.

Maşın tərcüməsi ciddi inkişaf yolu keçsə də, heç bir mövcud tərcümə sistemi hələ də insan qədər mükəmməl tərcümə edə bilmir. Keyfiyyətli maşın tərcüməsi prosesində ən böyük maneələrdən biri ikimənalı, yaxud çoxmənalı sözün tərcüməsinin düzgün şəkildə aparılmasıdır. Maşın tərcüməsində bu cür nöqsanların həlli ikimənalı və ya çoxmənalı söz və ya sözlərlə əhatə olunan kontekstin təhlilindən asılıdır. Maşının daxilindəki mətndən bir söz çıxarılaraq tərcüməyə verildikdə tərcüməçi həmin mətnin kontekstindən xəbərsizdirse, o halda həmin sözün mənasını maşın tərcüməsi qədər dəqiq müəyyən edə bilməz, çünki maşına yerləşdirilmiş mətndə həmin sözün mənalarını program ayırd edə bilir. Program düzgün qurul-

muşdursa, mətnindəki həmin sözün mənalarını da düzgün müəyyənləşdirir.

Maşın tərcüməsində semantik təhlilin əsas məqsədi ikimənalılıq və çoxmənalılıq problemini həll etməkdir. Bundan başqa, semantik təhlilin məqsədlərindən biri də xüsusi sahəyə aid bütün leksik vahidləri toplamaq və mənalarına görə qruplaşdırmaqdır. Burada hər bir elementin başqa mövcud elementlə və ya elementlərlə ayrı-ayrılıqlıda, həmçinin ümumilikdə əlaqəsini göstərmək vacibdir (I cədvəl)

I cədvəl

Sözlərin mənalarına görə qruplaşdırılması

		yaxın təməslə acılılı halda	gözləri geniş acılılıs təccüb içində
baxmaq	+		+
görmək	+		
nəzər salmaq	+		+
Göz gəzdirmək	+		+
müşahidə etmək	+		
gözdən keçirmək	+		
gözünü zilləmək	+		+
maritlamaq	+		
izləmək	+		
süzmək	+		
ağzını ayırmaq	+		

Mətndə iki leksem mənasının üst-üstə düşdüyü halların, müəyyən leksik boşluqların olması da mətnin məna keyfiyyətinə o qədər xələl gətirmir, lakin belə boşluqlar, uygununsuzluqlar artdıqca mətn tərcümə və təhlil zamanı öz mənasını itirir. Ona görə də mətnin semantik təhlilində onomimlik və çoxmənalılıq problemləri həll olunmalıdır. Dildə leksem mənası dil sistemində semantik baxımdan bir-biri ilə əlaqəli olan başqa leksemərin mövcudluğundan asılıdır. Semantik təhlil nəzəriyyəsi dilin leksik vahidlərinin semantik cəhətdən əlaqəli olan sahələr üzrə təsnif olunmasını və onların semantik qruplara bölünməsini nəzərdə tutur.

İstər morfoloji, istər sintaktik səviyyələrdə “həll olunmamış problemlər” də maşın tərcüməsinin semantik təhlil mərhələsində araşdırılır (Mahmudov, 2013: 297).

Çoxmənalılıq problemi çoxmənalı sözün mənasının kontekstdə müəyyənləşdirilməsini və yerində işlənməsini həyata keçirən program qurmaqla həll edilə bilər. Bir sözlə, tərcümə programı sözün kontekstual mənasını sözlər haqqında insan biliyi ilə müqayisə ediləcək qədər məlumatə sahib olduğu hallarda düzgün seçə bilər.

İSTİFADƏ OLUNMUŞ ƏDƏBİYYAT

1. Hearst M.A. (1992). Automatic acquisition of hyponyms from large text corpora // Processing 14th International Conference on Computational Linguistics. Nantes, France: August 23-28.
2. Izutsu T. (1975). Kur'an'da Allah ve İnsan. Ankara: Ankara Üniversitesi Basımıvi.
3. Kazimov Q.Ş. (2008) .Seçilmiş əsərləri. 10 cilddə. I c. Bakı: Nurlan.
4. Lyons J. (1968). Introduction to Theoretical Linguistics. Cambridge: Cambridge University Press.
5. Mahmudov M.Ə. (2013). Kompüter dilçiliyi. Bakı: Elm və təhsil.
6. Vəliyeva K.A. (1996). Mətnin avtomatik sintaktik təhlili və sintezi. Bakı: Örnək.
7. Vəliyeva K.A., Məmmədova M. (2003). Mətnlərin avtomatik redaktəsi. Bakı: Azərbaycan Universiteti nəşriyyatı
8. Городецкий Б.Ю. (1976). Семантические проблемы построения автоматизированных систем обработки текстовой информации // Вычислительная лингвистика. М.: Наука.
9. Махмудов М.А. (1982). Разработка системы формально-морфологического анализа тюркской словоформы (на материале азербайджанского языка). Дис. ... канд. филол. наук. Баку: Элм.
10. Мочалова А.В. (2014). Алгоритм семантического анализа текста, основанный на базовых семантических шаблонах с удалением // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. № 5 (93).
11. Сокирко А. (2001). Семантические словари в автоматической обработке текста (по материалам системы ДИАЛИНГ). Дисс. ... канд. тех. наук. Москва.
12. aot.ru/docs/sokirko/sokirko-candid-4.html, (15.03.2020)
13. www.wikipedia.az, (04.02.2020)

*Камиля Велиева (Азербайджан)
Гульнара Джасфарова (Азербайджан)*

Семантический анализ слова в машинном переводе

Резюме

В лингвистике семантика изучает значение или согласование языковых компонентов (слов, фраз) в концепте человеческого сознания. В машинном переводе основным принципом семантического анализа является выявление значения слова. Кроме того, здесь решаются различные практические и теоретические проблемы семантического анализа, такие как вопрос-ответ и извлечение информации.

На этапе семантического анализа при машинном переводе исходя из значения слова готовятся формальные семантические модели. Эти модели стимулируют синтаксический и семантический анализ слова и процесс семантического анализа. В статье исследуются вопросы семантического анализа слова в машинном переводе. Авторы показывают возможность решения проблемы многозначности в машинном переводе путём создания программы, осуществляющей выявление значения многозначного слова лишь в контексте и уместное его использование.

Ключевые слова: *машинный перевод, семантический анализ, слово, значение, модель, словарь.*

*Kamila Valiyeva (Azerbaijan)
Gulnara Jafarova (Azerbaijan)*

Semantic analysis of the word in the machine translation

Abstract

In linguistics, semantics learns the meaning or the conforming of the language components (words, phrases) into each other in the concept of human thought. In the machine translation, the main principle of the semantic analysis is to learn determining the meaning. Besides, different practical and theoretical problems as question-answering system and revealing information belong to semantic analysis are disclosed here.

Formal semantic models are made from the meaning of the words in the level of semantic analysis of the machine translation. These models boost the syntactic and semantic analysis of the word and help the process of the semantic analysis. The issues of the semantic analysis of the word in the machine translation are researched in the article. The author shows the solution of the problem of polysemy in the machine translation by the creation of the program realizing the definition of the meaning of word in the context only and the usage in the relevant place.

Key words: *machine translation, semantic analysis, word, meaning, model, dictionary.*